FLORE
Repository istituzionale dell'Università degli Studi di Firenze

Decomposition Algorithms for Generalized Potential Games

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

Original Citation:

Decomposition Algorithms for Generalized Potential Games / F. Facchinei; V. Piccialli; M. Sciandrone. - In: COMPUTATIONAL OPTIMIZATION AND APPLICATIONS. - ISSN 0926-6003. - STAMPA. - 50:(2011), pp. 237-262.

Availability:

This version is available at: 2158/642114 since:

Terms of use:

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf)

Publisher copyright claim:

(Article begins on next page)

05 May 2023

Comput Optim Appl (2011) 50:237–262
DOI 10.1007/s10589-010-9331-9

Decomposition algorithms for generalized potential games

Francisco Facchinei · Veronica Piccialli · Marco Sciandrone

Received: 23 November 2009 / Published online: 8 May 2010
© Springer Science+Business Media, LLC 2010

Abstract We analyze some new decomposition schemes for the solution of generalized Nash equilibrium problems. We prove convergence for a particular class of generalized potential games that includes some interesting engineering problems. We show that some versions of our algorithms can deal also with problems lacking any convexity and consider separately the case of two players for which stronger results can be obtained.

Keywords Generalized Nash equilibrium problem · Generalized potential game · Decomposition · Regularization

1 Introduction

In this paper we consider decomposition algorithms for the solution of a Generalized Nash Equilibrium Problem (GNEP for short). The GNEP extends the classical Nash Equilibrium Problem (NEP) by assuming that each player's feasible set can depend

To Liqun Qi on the occasion of his 65th birthday, with friendship and admiration.

F. Facchinei (✉)
Dipartimento di Informatica e Sistemistica Antonio Ruberti, Sapienza Università di Roma, Via Ariosto 25, 00185 Roma, Italy
e-mail: facchinei@dis.uniroma1.it

V. Piccialli
Dipartimento di Ingegneria dell'Impresa, Università di Tor Vergata, Via del Politecnico 1, 00133 Roma, Italy
e-mail: piccialli@disp.uniroma2.it

M. Sciandrone
Dipartimento di Sistemi e Informatica, Università di Firenze, Via S. Marta 3, 50139 Firenze, Italy
e-mail: sciandro@dsi.unifi.it

on the rival players' strategies. There are N players, and each player ν controls the variables $x^\nu \in \mathbb{R}^{n_\nu}$. We denote by $\boldsymbol{x}$ the vector formed by all these decision variables

$$\boldsymbol{x} \equiv \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix},$$

which has dimension $n := \sum_{\nu=1}^N n_\nu$ and by $\boldsymbol{x}^{-\nu}$ the vector formed by all the players' decision variables except those of player ν. To emphasize the ν-th player's variables within $\boldsymbol{x}$ we sometimes write $(x^\nu, \boldsymbol{x}^{-\nu})$ instead of $\boldsymbol{x}$.

Each player's strategy must belong to a set $X_\nu(\boldsymbol{x}^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$ that depends on the rival players' strategies. The aim of player ν, given the other players' strategies $\boldsymbol{x}^{-\nu}$, is to choose a strategy x^ν that solves the minimization problem

$$\text{minimize}_{x^\nu}\, \theta_\nu(x^\nu, \boldsymbol{x}^{-\nu}) \quad \text{subject to} \quad x^\nu \in X_\nu(\boldsymbol{x}^{-\nu}). \tag{1}$$

The GNEP is the problem of finding a vector $\bar{\boldsymbol{x}}$ such that each player's strategy $\bar{x}^\nu$ solves the player problem (given $\bar{\boldsymbol{x}}^{-\nu}$):

$$\theta_\nu(\bar{x}^\nu, \bar{\boldsymbol{x}}^{-\nu}) \le \theta_\nu(y^\nu, \bar{\boldsymbol{x}}^{-\nu}), \quad \forall y^\nu \in X_\nu(\bar{\boldsymbol{x}}^{-\nu}).$$

Such a point $\bar{\boldsymbol{x}}$ is called a (generalized) Nash equilibrium or, more simply, a solution of the GNEP.

We refer the interested reader to [11–13] for a state-of-the-art discussion on GNEPs. The GNEP has many important applications [11, 13, 18] and in the past ten years its use has spread from traditional settings in economy to many innovative models in engineering. Actually it is safe to say that a new branch of game theory has recently emerged that, using the words of Nobel Laureate Robert Aumann, we call *Game Engineering* and that is also widely known, especially in the computer science community, as *Algorithmic Game Theory* [18]. In this new setting the computability and the actual computation of equilibria plays a central role. Unfortunately, algorithms for the solution of general GNEPs are exceptionally difficult to analyze and useful convergence results are extremely scarce. In fact, while many algorithms have been proposed, especially in recent years, the conditions that guarantee convergence of the methods are very strong, technical, difficult to verify and, with the exception of "jointly convex problems" [4, 11, 13, 16, 22], in most cases it is not possible to prove convergence for a clearly defined and easily identifiable class of problems [11].

Decomposition algorithms are very natural and immediately spring to mind when considering GNEPs. Since in a Nash game every player is trying to minimize his own objective function, a natural approach is to consider an iterative algorithm based on a Gauss-Seidel scheme where at each iteration every player, given the strategies of the others, updates his own strategy by solving his optimization problem (1). The resulting algorithm is described below.

In Step 0, by "feasible" we mean a point that satisfies the constraints of all players.

Algorithms as the one just described are of the utmost importance because in most applications, in particular in the engineering field (see also Sect. 3), they describe the "behavior" of the players. Therefore, it is exactly the convergence of such type of

Algorithm 1 Gauss-Seidel best-response algorithm

`(S.0):` Choose a feasible starting point $\boldsymbol{x}_0 = (x_0^1, \ldots, x_0^N)$, and set $k := 0$.
`(S.1):` If $\boldsymbol{x}_k$ satisfies a suitable termination criterion: `STOP`.
`(S.2):` `for` $\nu = 1, \ldots, N$, compute a solution x_{k+1}^ν of

$$\begin{aligned} \min_{x^\nu} \quad & \theta_\nu(x_{k+1}^1, \ldots, x_{k+1}^{\nu-1}, x^\nu, x_k^{\nu+1}, \ldots, x_k^N) \\ \text{s.t.} \quad & x^\nu \in X_\nu(x_{k+1}^1, \ldots, x_{k+1}^{\nu-1}, x_k^{\nu+1}, \ldots, x_k^N). \end{aligned} \tag{2}$$

`end`
`(S.3):` Set $\boldsymbol{x}_{k+1} := (x_{k+1}^1, \ldots, x_{k+1}^N)$, $k \leftarrow k+1$, and go to `(S.1)`.

algorithms to a Nash equilibrium that actually justifies the use of Nash equilibria as a suitable solution concept.

Unfortunately, convergence of Algorithm 1 is hard to prove, see e.g. [11, 13]. Actually, we will show in Sect. 3 that one cannot expect to obtain convergence of Algorithm 1 even for rather well behaved problems. To overcome these difficulties we analyze a regularized version of Algorithm 1 inspired by the analysis in [15]. This regularized version essentially differs from Algorithm 1 only in Step 2, where a regularization term is added to the objective function (see (3) below).

Algorithm 2 Regularized Gauss-Seidel best-response algorithm

`(S.0):` Choose a feasible starting point $\boldsymbol{x}_0 = (x_0^1, \ldots, x_0^N)$, a positive regularization parameter $\tau_0 > 0$ and set $k := 0$.
`(S.1):` If $\boldsymbol{x}_k$ satisfies a suitable termination criterion: `STOP`.
`(S.2):` `for` $\nu = 1, \ldots, N$, compute a solution x_{k+1}^ν of

$$\begin{aligned} \min_{x^\nu} \quad & \theta_\nu(x_{k+1}^1, \ldots, x_{k+1}^{\nu-1}, x^\nu, x_k^{\nu+1}, \ldots, x_k^N) + \tau_k \|x^\nu - x_k^\nu\|^2 \\ \text{s.t.} \quad & x^\nu \in X_\nu(x_{k+1}^1, \ldots, x_{k+1}^{\nu-1}, x_k^{\nu+1}, \ldots, x_k^N). \end{aligned} \tag{3}$$

`end`
`(S.3):` Update τ_k. Set $\boldsymbol{x}_{k+1} := (x_{k+1}^1, \ldots, x_{k+1}^N)$, $k \leftarrow k+1$, and go to `(S.1)`.

In Sect. 2 we define a class of *Generalized Potential Games* (GPGs), expanding the well-known class of NEP potential games introduced in [17], for which convergence of Algorithm 2 can be established. This class of generalized potential games is certainly peculiar, but many problems of interest are actually GPGs, see [17] and Sect. 3 where we report two telecommunication applications that are GPGs; furthermore, it is easy to decide whether a problem is a GPG. We note from the outset that in principle a Nash equilibrium of a generalized potential game could be calculated by solving in a centralized way a single optimization problem. However, the solution of this centralized problem could be impossible in many situations both because it could turn out to be too difficult and because in many practical settings the application of a centralized algorithm is simply not possible, see Sect. 3 for further comments. In Sect. 3 we will also discuss more in detail the fact that, in spite of the possible optimization reformulation of a GPG, convergence of Algorithm 2 (or of any other practical distributed algorithm) cannot be derived by existing decomposition methods for optimization problems. In Sects. 4, 5, and 6 we consider three instances of Algorithm 2 under different assumptions. In particular, in Sect. 4, under convexity assumptions, we show convergence for a fixed value of the regularization parameters τ_k. In Sect. 5 we drop the convexity assumptions and prove convergence by introducing an updating rule for τ_k that drives it to zero. Finally, in Sect. 6 we consider the case of two players, where it is possible to show convergence of Algorithm 1 without regularization (i.e. convergence of Algorithm 2 with $\tau_k = 0$) and for a broader class of GPGs.

Summarizing, the main contribution of this paper is a convergence theory for Algorithm 2 that is applicable to a class of easily identifiable problems with interesting engineering applications. Generalized potential games constitute the first class of GNEPs that can be solved by distributed algorithms and overall our results extend considerably the range of solvable GNEPs.

2 Generalized potential games

Roughly speaking a generalized potential game is a GNEP where the players are (unknowingly) minimizing the same function and where the feasible set of each player is the "section" of a larger set in the product space $\mathbb{R}^n$. Formally we have the following definition.

Definition 2.1 A GNEP is a Generalized Potential Game if:

(a) There exists a nonempty, closed set $\boldsymbol{X} \subseteq \mathbb{R}^n$ such that, for all $\nu = 1, \ldots, N$,

$$X_\nu(\boldsymbol{x}^{-\nu}) \equiv \{x^\nu \in D_\nu : (x^\nu, \boldsymbol{x}^{-\nu}) \in \boldsymbol{X}\}, \tag{4}$$

where $D_\nu \subseteq \mathbb{R}^{n_\nu}$ are nonempty, closed sets such that $\prod_{\nu=1}^N D_\nu \cap \boldsymbol{X} \neq \emptyset$ (i.e. the "feasible set" of the game is non empty).

(b) There exists a *continuous* function $P(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$ such that for all ν, for all $\boldsymbol{x}^{-\nu}$ (such that $X_\nu(\boldsymbol{x}^{-\nu})$ is not empty), and for all $y^\nu, z^\nu \in X_\nu(\boldsymbol{x}^{-\nu})$

$$\theta_\nu(y^\nu, \boldsymbol{x}^{-\nu}) - \theta_\nu(z^\nu, \boldsymbol{x}^{-\nu}) > 0$$

implies

$$P(y^\nu, \boldsymbol{x}^{-\nu}) - P(z^\nu, \boldsymbol{x}^{-\nu}) \geq \sigma(\theta_\nu(y^\nu, \boldsymbol{x}^{-\nu}) - \theta_\nu(z^\nu, \boldsymbol{x}^{-\nu})), \tag{5}$$

where $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ is a forcing function: $\lim_{k\to\infty} \sigma(t_k) = 0 \Rightarrow \lim_{k\to\infty} t_k = 0$.

Our definition is inspired by the classical definition of Monderer and Shapley [17]. In fact, if (i) the condition (b) becomes that a function P exists such that (5) holds

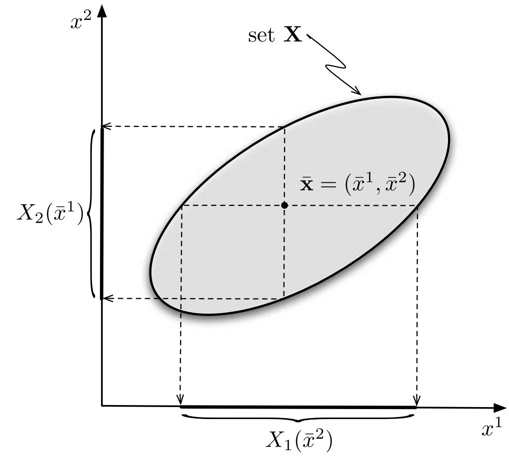

Fig. 1 Condition (a): The set $\boldsymbol{X}$ is convex and so are the sets X_1 and X_2

with equality and with the forcing function equal to the identity and, most importantly, (ii) the "big set" $\boldsymbol{X}$ is given by the cartesian product of lower dimensional sets, $\boldsymbol{X} = X_1 \times \cdots \times X_N$, with $X_\nu \subseteq \mathbb{R}^{n_\nu}$, then the GPG is just a(n exact) potential NEP as defined in [17]. We generalize this classical definition of potential game in several directions and we believe the following points are important.

1. We do not assume a cartesian product structure for $\boldsymbol{X}$, so that we are really dealing with a GNEP and not with a NEP;
2. We do not make any of the usual convexity assumptions either on the set $\boldsymbol{X}$ or on the functions $\theta_\nu(\cdot, \boldsymbol{x}^\nu)$.

As a side, more technical point, we underline that in our definition we also have the freedom to choose a forcing function σ different from the identity. This might be important in some cases. But the really important departure from the classical definition of potential game is in the presence of coupling constraints that are represented by the set $\boldsymbol{X}$. This feature permits to deal with a host of new, interesting problems, as exemplified by the applications described in the next section.

Condition (a) in the definition of GNEP is obviously related to the class of "jointly convex" GNEPs, also known as GNEPs with shared constraints, see [11, 13, 22], where however the set $\boldsymbol{X}$ must be convex. Indeed, our definition of GPG borrows from both the classical theory of potential games and the theory of games with shared constraints.

Condition (a) is illustrated geometrically in Figs. 1, 2, and 3, where for simplicity we took the sets D_ν to be equal to the whole space $\mathbb{R}^{n_\nu}$.

Figure 1 represents the simplest case: both the set $\boldsymbol{X}$ and the feasible sets $X_\nu(\boldsymbol{x}^{-\nu})$ are convex. With regard to the feasible sets, this is essentially the setting of jointly convex problems. In Fig. 2 we have a more general situation in which the set $\boldsymbol{X}$ is not convex, although the feasible sets $X_\nu(\boldsymbol{x}^{-\nu})$ are convex so that, assuming the players' objective functions are also convex, the players' optimization problems are still convex. As far as we are aware of, this setting is already new in the literature. Finally, Fig. 3 illustrates the hardest case: both the set $\boldsymbol{X}$ and (some of) the sets $X_\nu(\boldsymbol{x}^{-\nu})$ are

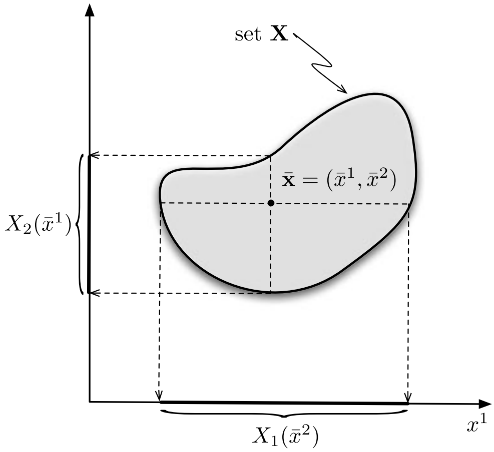

Fig. 2 Condition (a): The set X is non-convex, but the sets X_1 and X_2 are convex

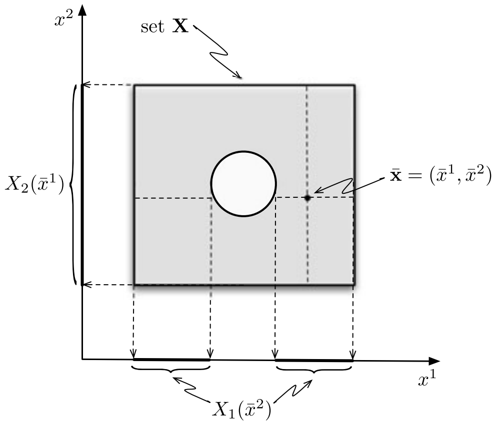

Fig. 3 Condition (a): The set X is non-convex, and the sets X_1 and X_2 might be non-convex

not convex. In this case, regardless of the objective functions, the players' problems are not convex. We will show that suitable variants of our basic Algorithm 2 can tackle all these cases.

Suppose now that, as usual, the set $\boldsymbol{X}$ is defined by some constraints

$$\boldsymbol{X} \equiv \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) \leq 0\},$$

where $g : \mathbb{R}^n \to \mathbb{R}^m$ is continuous. Then it is easy to see that condition (a) is equivalent to saying that these constraints $g(\boldsymbol{x}) \leq 0$ are shared by all players

$$X_\nu(\boldsymbol{x}^{-\nu}) \equiv \{x^\nu \in D_\nu : g(x^\nu, \boldsymbol{x}^{-\nu}) \leq 0\}. \tag{6}$$

The more favorable case of a convex $\boldsymbol{X}$ then corresponds to the fact that the D_ν are convex and the g_i are convex with respect to all variables $\boldsymbol{x}$ (from which the name jointly convex comes from).

We now examine condition (b) in Definition 2.1. This condition says that there is a single function P that, in some sense, reflects the changes in the players' objective functions. It can be easily checked that a *global solution* of the following optimization problem

$$\text{minimize}_{\boldsymbol{x}}\ P(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{x} \in \boldsymbol{X}, \quad x^\nu \in D_\nu,\ \nu = 1, \ldots, N \tag{7}$$

yields a Nash equilibrium. This shows that GPGs provide a bridge between game theory and optimization and obviously suggests a possible avenue for the solution of the game: the solution of optimization problem (7). However, the solution of problem (7) might be not simple since this is, in general, a global optimization problem. In this sense, decomposition methods have a clear advantage in that each player's subproblem could be easy, even when (7) is difficult. Furthermore, in many practical situations, the solution of problem (7) can not be conceptually considered (see next section). We observe that there is a large body of literature related to decomposition methods for the solution of optimization problems. However, we are not aware of any result that could be applied to (7) in order to calculate a (global minimum and therefore a) Nash equilibrium unless very stringent assumptions are made both on the function P and the set $\boldsymbol{X}$.

Verification of condition (b) in Definition 2.1 is usually rather straightforward. In some cases of interest the objective functions do not depend on the other players' variables, $\theta_\nu(\boldsymbol{x}) = \theta_\nu(x^\nu)$, so that the interaction of the players takes places only at the level of feasible sets. In this event it is immediate to see that condition (b) is satisfied with the potential function P simply given by the sum of the objective functions of all players. Another common case is when $\theta_\nu(\boldsymbol{x}) = c(\boldsymbol{x}) + d_\nu(x^\nu)$, that is when the objective functions have a common term c which is the same for all players plus an additional cost related only to x^ν. Also in this case, it is immediate to verify that condition (b) is satisfied with $P(\boldsymbol{x}) = c(\boldsymbol{x}) + \sum_{\nu=1}^N d_\nu(x^\nu)$.

The main contribution of this paper is the development of decomposition algorithms for the calculation of Nash equilibria of a GPG, under the following weak conditions, that we assume to hold throughout.

A1 For every $\nu = 1, \ldots, N$ the objective function $\theta_\nu(\boldsymbol{x})$ is continuous on $\boldsymbol{X}$.
A2 The feasible sets $X_\nu(\cdot)$ are inner-semicontinuous relative to $\text{dom}(X_\nu)$.[1]

The inner-semicontinuity requirement says that if $\bar{\boldsymbol{x}}$ belongs to $\boldsymbol{X}$ and, for any player ν, we consider any sequence $\{\boldsymbol{x}_k^{-\nu}\} \subset \text{dom}(X_\nu)$ such that $\{\boldsymbol{x}_k^{-\nu}\} \to \bar{\boldsymbol{x}}^{-\nu}$, then we may find points $x_k^\nu \in X_\nu(\boldsymbol{x}_k^{-\nu})$ such that $\{x_k^\nu\} \to \bar{x}^\nu$. This is not too restrictive a requirement. For example it is certainly satisfied if $\boldsymbol{X}$ is polyhedral, see [21, Example 9.35] and at any point in the interior of $\text{dom}(X_\nu)$, provided that $\boldsymbol{X}$ is convex, see [21, Theorem 5.9(b)]. We further note that for any of the sets in Figs. 1–3, the inner-semicontinuity property holds. Inner-semicontinuity may be more difficult to check on the boundary of $\text{dom}(X_\nu)$, even when the set $\boldsymbol{X}$ is convex. In this latter case some suitable (Slater type) constraint qualification may be needed. For details we refer the interested reader to [21] and, for a more in-depth treatment, to Chap. 3 of [5].

[1]We recall that $\text{dom}(X_\nu)$ is the set of points $\boldsymbol{x}^{-\nu}$ for which $X_\nu(\boldsymbol{x}^{-\nu})$ is non empty.

We remark that by their definition, and by the closure of all sets involved, the mappings $X_\nu(\cdot)$ are also closed, i.e. if we have two sequences $\{x_k^\nu\} \to \bar{x}^\nu$ and $\{\boldsymbol{x}_k^{-\nu}\} \to \bar{\boldsymbol{x}}^{-\nu}$ such that $x_k^\nu \in X_\nu(\boldsymbol{x}_k^{-\nu})$, then $\bar{x}^\nu \in X_\nu(\bar{\boldsymbol{x}}^{-\nu})$. This fact will be used without further notice throughout the paper.

In the setting above we will consider three distinct cases: for every ν and for every fixed $\boldsymbol{x}^{-\nu}$ the players' problems (1) are convex (Sect. 4); the case in which the same problems are not necessarily convex (Sect. 5); the case in which there are only two players (Sect. 6). We remark that while we guarantee that our algorithms converge to Nash equilibria of the GPG, these Nash equilibria are not necessarily the solutions of (7).

3 Discussion and examples

This section is devoted to the discussion and clarification through examples of some of the issues discussed in the previous two sections.

3.1 Algorithm 1 does not work

We begin with an example showing that, in contrast with what stated for example in [10], Algorithm 1 in general does not converge to a Nash equilibrium of a potential game, even under favorable assumptions. To show this we consider a (slight) variant of a classical counterexample of Powell [20].

Consider a potential Nash game with 3 players. Each player controls one variable (which we denote by u, v, z, having set for simplicity, $x_1^1 = u, x_1^2 = v, x_1^3 = z$). The three players minimize the same objective function $\theta_1 = \theta_2 = \theta_3 = P$ with $P : R^3 \to R$ given by:

$$P(\boldsymbol{x}) = -uv - vz - uz + (u-1)_+^2 + (-u-1)_+^2 \\ + (v-1)_+^2 + (-v-1)_+^2 + (z-1)_+^2 + (-z-1)_+^2,$$

where $(g)_+ = \max\{0, g\}$. Assume further that $X_1 = X_2 = X_3 = [-10, 10]$. We see that P is a nonconvex function although it is component-wise convex and the resulting game is obviously a GPG, actually even a potential game according to [17].

Suppose that the three players play in the order 1, 2 and 3. This is equivalent, in the jargon of decomposition optimization methods, to applying the Gauss-Seidel decomposition method to the minimization of P over $X_1 \times X_2 \times X_3$. Powell [20] showed that if the starting point $\boldsymbol{x}_0 = (u_0, v_0, z_0)$ is the point $(-1-\epsilon, 1+\frac{1}{2}\epsilon, -1-\frac{1}{4}\epsilon)$ (we take $\epsilon \in (0, 8]$ to keep feasibility) the steps of the GS method produce the following sequence of points in the first 6 iterations (Powell showed that these points are the unconstrained minimizers of each player's subproblem, therefore they are *a fortiori* the minimizers over the set $[-10, 10]$)

$$\begin{pmatrix} 1+\frac{1}{8}\epsilon \\ 1+\frac{1}{2}\epsilon, \\ -1-\frac{1}{4}\epsilon \end{pmatrix} \to \begin{pmatrix} 1+\frac{1}{8}\epsilon \\ -1-\frac{1}{16}\epsilon \\ -1-\frac{1}{4}\epsilon \end{pmatrix} \to \begin{pmatrix} 1+\frac{1}{8}\epsilon \\ -1-\frac{1}{16}\epsilon \\ 1+\frac{1}{32}\epsilon \end{pmatrix} \to \begin{pmatrix} -1-\frac{1}{64}\epsilon \\ -1-\frac{1}{16}\epsilon \\ 1+\frac{1}{32}\epsilon \end{pmatrix}$$

$$\to \begin{pmatrix} -1-\frac{1}{64}\epsilon \\ 1+\frac{1}{128}\epsilon \\ 1+\frac{1}{32}\epsilon \end{pmatrix} \to \begin{pmatrix} -1-\frac{1}{64}\epsilon \\ 1+\frac{1}{128}\epsilon \\ -1-\frac{1}{256}\epsilon \end{pmatrix}.$$

This last point is the same as the starting point $\boldsymbol{x}_0$ except that ϵ has been replaced by $\frac{1}{64}\epsilon$. Therefore the calculated sequence of points has six limit points given by

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$$

None of these 6 limit points is a Nash equilibrium, since it is easy to check that in each of these points, at least one component of the gradient of P is nonzero and no constraint is active. So we see that in a problem that is actually a NEP (as opposed to GNEP), and in which every player's subproblem is convex in the player's variables and has a compact feasible set, Algorithm 1 fails completely.

The example we just examined also helps understanding another point we made: while the minimization of P over the rectangle $X_1 \times X_2 \times X_3$ is in this case a difficult non convex optimization problem, the subproblems each player should solve according to Algorithm 2 are (very simple) strongly convex optimization problems in one variable.

3.2 Any minimization algorithm for the solution of (7) is of little help

One could think that the failure analyzed in the previous subsection is due to the particular (decomposition) algorithm we chose to solve (7). Here we show that, whatever the algorithm (be it centralized or decomposition-type), we cannot hope to practically solve a general GPG by solving the minimization problem (7). To illustrate this point easily, we consider a modification of the problem in the previous subsection, where we only change the set $\boldsymbol{X}$ by setting $\boldsymbol{X} = [-10, 10] \times [-10, 10] \times [-10, 10] \cap \{\boldsymbol{x} \in \mathbb{R}^3 : z^2 - 1 \geq 0\}$. Thus we simply intersected the set $\boldsymbol{X}$ in the previous subsection with the set $\{\boldsymbol{x} \in \mathbb{R}^3 : z^2 - 1 \geq 0\}$. In practice instead of a cube centered in the origin with sides' length of 20, we are considering two disjoint "slices" of this cube, having eliminated all the points with $z \in (-1, 1)$. Consider now the feasible point $\bar{\boldsymbol{x}} = (1, 1, -1)$. In this point we have $\nabla P(\bar{\boldsymbol{x}}) = (0, 0, -2)$ and only the constraint $z^2 - 1 \geq 0$ is active. It is easy to check, taking the multiplier relative to this constraint equal to 1 and the remaining ones equal to 0, that $\bar{\boldsymbol{x}}$ is stationary for the minimization problem (7). However $\bar{\boldsymbol{x}}$ is not a Nash equilibrium, since $P(1, 1, -1) = 1$, while $P(1, 1, 1) = -3$. Since $(1, 1, 1)$ is feasible we see that the third player can improve the objective function by unilaterally deviating from $\bar{\boldsymbol{x}}$, thus showing that $\bar{\boldsymbol{x}}$ is not a Nash equilibrium point.

This example shows us that in general it might not be a good idea to try to solve the minimization problem (7) to compute an equilibrium, since standard methods will generate stationary points and not global solutions. Obviously, if problem (7) is convex, any stationary point will be a global solution and so the approach becomes feasible. But even in the favorable condition in which (7) is convex, in some situations the solution of (7) in a centralized way may still be not feasible; this is one of the topics of the next two subsections (see, in particular, the two final paragraphs of these two subsections). From a different point of view, even if problem (7) is convex, in general, the set of Nash equilibria will be larger than the set of solutions of the optimization problem; so if we try to calculate a Nash equilibrium by solving the optimization problem (7) we are restricting our target and it is intuitively clear that this might make the computation more difficult.

3.3 Flow control in multi-hop communication networks

The aim of this subsection is twofold. On the one hand we want to illustrate an interesting application whose formulation gives rise to a GPG and, on the other hand, we use this application to show why solving (7) might be unfeasible in some practical settings. The model we discuss belongs to a rather well studied class of problems, see for example the review papers [1, 2]; here we follow in particular the setting of [23].

We consider a general transmission network model based on fluid approximation. The topology of the network is characterized by a set of nodes $\mathcal{V} = \{1, \ldots, V\}$ and a set of links $\mathcal{L} = \{1, \ldots, L\}$ connecting the nodes (we assume that the network is connected). There are N active users (players); each user ν is uniquely associated to a connection between the source node s_ν and the destination node d_ν through a path $\mathcal{L}_\nu$ (predetermined by a routing algorithm), where $\mathcal{L}_\nu$ is the subset of links that form the path of user ν. The information flow routed through the path $\mathcal{L}_\nu$ by user ν is denoted by x^ν and it holds $x^\nu \in D_\nu = \{x^\nu \in \mathbb{R} : 0 \le x^\nu \le x^\nu_{\max}\}$, where $x^\nu_{\max}$ is a physical or regulatory positive upper bound. Each link ℓ has a capacity c_ℓ. If we introduce the $L \times N$ routing matrix A, defined by $A_{\ell,\nu} = 1$ if $\ell \in \mathcal{L}_\nu$ and 0 otherwise, the capacity constraints can be expressed in vector form as $A\boldsymbol{x} \le c$, where $c = (c_\ell)_{\ell=1}^L$. Finally, we define the set $\boldsymbol{X}$ of shared constraints as $\boldsymbol{X} = \{\boldsymbol{x} \in \mathbb{R}^N : A\boldsymbol{x} \le c\}$.

We can associate with this setting a GPG where $\boldsymbol{X}$ and D_ν are the sets defined above and the payoff function (to be minimized) of each player ν is

$$\theta_\nu(\boldsymbol{x}) = \sum_{\ell:\ell\in\mathcal{L}_\nu} P_\ell(\boldsymbol{x}) - U_\nu(x^\nu), \tag{8}$$

which is taken as the difference of a pricing function (the sum of the costs relative to each link on the path $\mathcal{L}_\nu$) and a reward U_ν associated to the flow x^ν sent by the player. The first term in the payoff function can be interpreted as the price that each user pays for using the network resources. We assume that each P_ℓ depends only on the sum of the flows on that link (the traffic on that link): $P_\ell(\boldsymbol{x}) = P_\ell(\sum_{\mu:\ell\in\mathcal{L}_\mu} x^\mu)$, with P_ℓ a convex function defined on $[0, c_\ell]$. The utility function U_ν, instead, is assumed to be, according to standard economic conditions and elastic traffic models, a strictly concave function defined on $[0, x^\nu_{\max}]$. Several pricing and reward functions have been proposed in the literature that satisfy the above assumptions; typical examples are

$$P_\ell = \frac{b_\ell}{\varepsilon + c_\ell - \sum_{\mu:\ell\in L_\mu} x^\mu}, \quad \text{and} \quad U_\nu = a_\nu \log(1 + x^\nu),$$

where a_ℓ, b_ℓ and ε are given positive constants. Note that, under our assumptions, the objective functions $\theta_\nu(x^\nu, x^{-\nu})$ are strictly convex in x^ν for every fixed $x^{-\nu}$. Actually, since the composition of a convex function with a linear function of $\boldsymbol{x}$ is convex, we have that θ_ν is also convex in $\boldsymbol{x}$. Finally, we also assume that all the P_ℓ and U^ν are continuously differentiable. It is very easy to verify that the game is a GPG with potential function given by

$$P(\boldsymbol{x}) = \sum_{\ell\in\mathcal{L}} P_\ell\left(\sum_{\mu:\ell\in\mathcal{L}_\mu} x^\mu\right) - \sum_{\nu=1}^{N} U_\nu(x^\nu). \tag{9}$$

We could compute a solution of this GPG by solving the optimization problem (7). Note that in this case the situation is very favorable because P is convex with respect to $\boldsymbol{x}$ and the sets $\boldsymbol{X}$ and D_ν are also convex and therefore any standard *centralized* optimization algorithm will lead to a global solution. And yet it should be clear from the type of application that such a centralized algorithm cannot be practically implemented in most situations of interest (think to the internet), since this would require a high degree of coordination among selfish users and also an unbearable overload of information exchange. If one turns to optimization decomposition technique, one finds that the only case that can be practically dealt with by standard decomposition methods is the case in which the set has a cartesian product structure $\boldsymbol{X} = X_1 \times \cdots \times X_N$, i.e. the case of NEPs, which does not cover our setting. On the other hand it is clear that this telecommunication problem fully fits our framework.

3.4 Joint rate and power optimization in interference limited wireless communication networks

We complete this section with a short analysis of a further telecommunication application that can give rise to problems more complex than the one considered in the previous subsection. More specifically, we consider the joint optimization of transmit powers and information rates in wireless (flat-fading) communication networks. Our description is mostly based on the setting of [19]; see also [7] for a broader picture.

We consider an N-users scalar Gaussian interference channel. In this model, there are N transmitter/receiver pairs, where each transmitter wants to communicate with its corresponding receiver over a scalar channel, which may represent time or frequency domains, affected by Gaussian noise. This mathematical model is sufficiently general to encompass many communication systems of practical interest, such as peer-to-peer networks, wireless flat-fading ad-hoc networks, and Code Division Multiple Access (CDMA) single/multicell cellular systems [24]. Here, we assume without loss of generality that all pairs utilize a CDMA transmission scheme [25] and share the same bandwidth (thus in principle they may interfere with each other). We denote by $G_{\nu\nu} > 0$ the effective channel gain of link ν (including the multiplicative spreading gain, antenna gain and coding gain), while $G_{\nu\mu} \geq 0$ denotes the effective (cross-)channel gain between the transmitter μ and the receiver ν. The transmit power of each transmitter ν is denoted by p^ν. To avoid excessive signalling and the need of coordination among users, we assume that encoding/decoding on each link is performed independently of the other links and no interference cancellation techniques are used. Hence, multiuser interference is treated as additive noise at each receiver. Within this setup, under mild information theoretical assumptions (see, e.g., [9]), the maximum information rate R^ν achievable over link ν for a given transmit power profile $\mathbf{p} = (p^\nu)_{\nu=1}^N$ of the users is

$$R^\nu(\mathbf{p}) = \log(\mathrm{SIR}^\nu(\mathbf{p})) \geq 0 \quad \text{with } \mathrm{SIR}^\nu(\mathbf{p}) \triangleq \frac{G_{\nu\nu} p^\nu}{\sum_{\mu\neq\nu} G_{\nu\mu} p^\mu} \tag{10}$$

where in the definition of the signal-to-noise-plus-interference ratio SIR^ν receiver noise is assumed to be dominated by the multiuser interference and is neglected in this model. This assumption is satisfied by most practical interference limited systems [19, 26]. Moreover, in these systems we have $\mathrm{SIR} \gg 1$, since it represents the effective SIR after spreading gain, antenna gain, and coding gain (see, e.g., [25]). Note that, in view of these assumptions, the SIR is homogeneous in the transmitters' powers, and a scaling of the powers does not affect the SIR. Therefore, in the sequel we assume that the powers have been normalized so that $\sum_{\nu=1}^N p^\nu = 1$.

Within the setup above, the strategy of each player ν (user) is the transfer rate x^ν at which data are sent over the link and the transmit power p^ν that supports such a rate x^ν. A transfer rate profile $\boldsymbol{x} = (x^\nu)_{\nu=1}^N \in \mathbb{R}_+^N$ is feasible if it is possible for the system to simultaneously transfer data over the network at the specific rates $\mathbf{x}$ for some power vector $\mathbf{p}$. The rate-region of the system is the set of feasible transfer rates $\mathbf{x}$, which is formally defined as

$$\mathcal{R} \triangleq \{\boldsymbol{x} \in \mathbb{R}_+^N \mid \exists \mathbf{p} \geq 0 \text{ such that } x^\nu \leq R^\nu(\mathbf{p}), \forall \nu = 1, \ldots, N\} \tag{11}$$

where $R^\nu(\mathbf{p})$ is defined in (10). It is not difficult to show that the rate-region $\mathcal{R}$ is a convex set [19]. An equivalent characterization of the rate-region $\mathcal{R}$ useful for our purpose can be given through the so called Perron-Frobenius eigenvalue [6] of a non-negative matrix, as detailed next. We recall that a square matrix M is non-negative, $M \geq 0$, if all its elements are non-negative and that if M is also primitive (meaning that all elements of M^k are positive for some k) then M has a unique strictly positive (real) eigenvalue $\lambda(M)$ which is larger in modulus than any other eigenvalue of M. Define two matrices D and $\tilde{G}$:

$$D(\boldsymbol{x}) = \mathrm{diag}\left(\frac{e^{x^\nu}}{G_{\nu\nu}}\right), \quad \tilde{G}_{\nu\mu} = \begin{cases} G_{\nu\mu} & \text{if } \nu \neq \mu, \\ 0 & \text{otherwise.} \end{cases}$$

Using the above definition, the following equivalence can be shown to hold [19]

$$x^\nu \leq R^\nu(\mathbf{p}), \quad \forall \nu \quad \Leftrightarrow \quad \mathbf{p} \geq D(\boldsymbol{x})\tilde{G}\mathbf{p} \tag{12}$$

which, using the mentioned results on non-negative matrices and recalling the normalization on $\mathbf{p}$, leads to the following equivalent expression for the rate-region $\mathcal{R}$

$$\mathcal{R} = \{\boldsymbol{x} \in \mathbb{R}_+^N \mid \lambda(D(\boldsymbol{x})\tilde{G}) \leq 1\}. \tag{13}$$

It is important to note that it can be proved that the function $\boldsymbol{x} \mapsto \lambda(D(\boldsymbol{x})\tilde{G})$ is continuously differentiable and convex [19].

The goal of each player is then to maximize his utility function $U_\nu(x^\nu)$, which is assumed to depend only on its own transmission rate x^ν, subject to the rate constraints in (13). The form of this utility function can range from very simple to complex expressions and may take into account several system level metrics and Quality of Service requirements (see, e.g., [7]). In general this function could also be non concave; we are not interested here in its precise expression. Stated in mathematical terms, we have a game where each players' problem is

$$\begin{aligned} \max_{x^\nu} \quad & U_\nu(x^\nu), \\ & \lambda(D(\boldsymbol{x})\tilde{G}) \le 1, \\ & x^\nu \ge 0. \end{aligned} \tag{14}$$

Therefore, the problems (14) define a GPG with (possibly) non concave objective potential function $P(\boldsymbol{x}) = \sum_{\nu=1}^N U_\nu(x^\nu)$ and a convex set $\boldsymbol{X} = \mathcal{R}$. The power profile $\mathbf{p}$ supporting the optimal rates $\mathbf{x}^\star$ that are solutions of this game can be computed solving the linear system $\mathbf{p} \ge D(\mathbf{x}^\star)\tilde{G}\mathbf{p}$, $\sum_{\nu=1}^N p^\nu = 1$. This can be done using some of the (distributed) algorithms proposed in the wide literature of power control problems (e.g., [14, 26]).

Again we note that the application of the decomposition Algorithm 2 reduces the problem of finding a Nash equilibrium to that of finding a global minimum of a univariate function on an interval. On the other hand, even assuming that $P(\boldsymbol{x})$ is concave, which is not always the case, the solution of problem (7) is by no means trivial, both with centralized or decomposition methods.

4 The case of N players with convexity

This is the first of three sections where we examine the behavior of Algorithm 2 under different assumptions on the convexity of the players' subproblems and on their number. We recall for convenience that each player's problem is

$$\text{minimize}_{x^\nu}\ \theta_\nu(x^\nu, \boldsymbol{x}^{-\nu}) \quad \text{s.t.} \quad x^\nu \in X_\nu(\boldsymbol{x}^{-\nu}) = \{x^\nu \in D_\nu : (x^\nu, \boldsymbol{x}^{-\nu}) \in \boldsymbol{X}\}. \tag{15}$$

We say that a point $\boldsymbol{x}$ is feasible if $x^\nu \in X_\nu(\boldsymbol{x}^{-\nu})$ for every player ν.

We begin our analysis by considering the case where every player's subproblem is convex. Specifically, we suppose that the following assumption holds.

A3 The objective functions $\theta_\nu(\cdot, \boldsymbol{x}^{-\nu})$ and the feasible sets $X_\nu(\boldsymbol{x}^{-\nu})$ are convex.

Under this assumption (implied by all commonly made assumptions in the field) we will show that we can take a fixed τ in the general Algorithm 2 and prove convergence. Algorithm 2 thus becomes:

Note that the vector $\boldsymbol{x}_{k,\nu+1}$ defined in the algorithm has no specific role in the algorithm itself; we introduce it for reference purposes only. As usual in a Gauss-Seidel procedure, for each iteration k the algorithm solves consecutively N optimization subproblems, one for each player ν. Therefore, player $\bar{\nu}$ calculates the optimal

Algorithm 3 Regularized Gauss-Seidel—Convex Subproblems

`(S.0):` Choose any feasible starting point $\boldsymbol{x}_0 = (x_0^1, \dots, x_0^N)$, a positive regularization parameter $\tau > 0$ and set $k := 0$.

`(S.1):` If $\boldsymbol{x}_k$ satisfies a suitable termination criterion: `STOP`.
Otherwise set $\boldsymbol{x}_{k,1} = \boldsymbol{x}_k$.

`(S.2):` `for` $\nu = 1, \dots, N$, compute a solution x_{k+1}^ν of

$$\begin{array}{ll} \min_{x^\nu} & \theta_\nu(x_{k+1}^1, \dots, x_{k+1}^{\nu-1}, x^\nu, x_k^{\nu+1}, \dots, x_k^N) + \tau \|x^\nu - x_k^\nu\|^2 \\ \text{s.t.} & x^\nu \in X_\nu(x_{k+1}^1, \dots, x_{k+1}^{\nu-1}, x_k^{\nu+1}, \dots, x_k^N). \end{array} \tag{16}$$

Set $\boldsymbol{x}_{k,\nu+1} = (x_{k+1}^1, \dots, x_{k+1}^\nu, x_k^{\nu+1}, \dots, x_k^N)$.
`end`

`(S.3):` Set $\boldsymbol{x}_{k+1} := (x_{k+1}^1, \dots, x_{k+1}^N)$, $k \leftarrow k+1$, and go to `(S.1)`.

---

solution of his problem using the new information for all players $\nu < \bar{\nu}$ and the old information for all players $\nu > \bar{\nu}$. The vector $\boldsymbol{x}_{k,\bar{\nu}+1}$ thus collects the new points x_{k+1}^ν for each players $\nu \le \bar{\nu}$ and the old points x_k^ν for $\nu > \bar{\nu}$.

We note that Algorithm 3 is well defined, since all subproblems (16) always have a (unique) solution. This is easily seen because all these subproblems are strongly convex, thanks to the term $\tau \|x^\nu - x_k^\nu\|^2$ and Assumption **A3**.

We preliminarily show that all points generated by Algorithm 3, including the "intermediate" points $\boldsymbol{x}_{k,\nu}$, are feasible according to the definition given immediately after (15).

Lemma 4.1 *For every k and every ν, $\boldsymbol{x}_{k,\nu}$ is feasible.*

Proof Assume that $\boldsymbol{x}_{k,\nu}$ is feasible. We first show that $\boldsymbol{x}_{k,\nu+1}$ is feasible. By definition,

$$\boldsymbol{x}_{k,\nu} = (x_{k+1}^1, \dots, x_{k+1}^{\nu-1}, x_k^\nu, \dots, x_k^N),$$

$$\boldsymbol{x}_{k,\nu+1} = (x_{k+1}^1, \dots, x_{k+1}^\nu, x_k^{\nu+1}, \dots, x_k^N).$$

The feasibility of $\boldsymbol{x}_{k,\nu}$ implies that $x_{k+1}^i \in D_i$ for all $i \in \{1, \dots, \nu-1\}$, $x_k^j \in D_j$ for all $j \in \{\nu, \dots, N\}$ and that $(x_{k+1}^1, \dots, x_{k+1}^{\nu-1}, x_k^\nu, \dots, x_k^N) \in \boldsymbol{X}$. By definition of x_{k+1}^ν we have that $x_{k+1}^\nu \in X_\nu(\boldsymbol{x}_{k,\nu}^{-\nu})$, that is $x_{k+1}^\nu \in D_\nu$, and $(x_{k+1}^1, \dots, x_{k+1}^\nu, x_k^{\nu+1}, \dots, x_k^N) \in \boldsymbol{X}$, namely $\boldsymbol{x}_{k,\nu+1}$ is feasible. This fact, together with $\boldsymbol{x}_0 \in X_1(\boldsymbol{x}_0^{-1}) \times \cdots \times X_N(\boldsymbol{x}_0^{-N})$, $\boldsymbol{x}_{k,1} = \boldsymbol{x}_k$, and $\boldsymbol{x}_{k,N+1} = \boldsymbol{x}_{k+1}$, completes the proof. □

Before proving the main convergence theorem of this section we still need a technical result about the behavior of subgradients of functions of two (groups of) variables. This result is needed in order to avoid differentiability assumptions on the objective functions θ_ν. Essentially, the proposition below states that a certain partial gradient is a locally bounded, closed map.

Proposition 4.2 *Let $f : \mathbb{R}^s \times \mathbb{R}^t \to \mathbb{R}$ be given and assume that f is locally Lipschitz continuous around a point $(\bar{u}, \bar{v}) \in \mathbb{R}^s \times \mathbb{R}^t$ and such that $f(\cdot, v)$ is convex for every v in a neighborhood of $\bar{v}$. Let $\{(u_k, v_k)\}$ be a sequence of points converging to $(\bar{u}, \bar{v})$ and let $\{\xi_k\}$, with $\xi_k \in \partial_u f(u_k, v_k)$ be a sequence of (Clarke's) partial generalized gradients. Then, every limit point $\bar{\xi}$ of this sequence (and there is at least one such limit point) belongs to $\partial_u f(\bar{u}, \bar{v})$.*

Proof By [8, Proposition 2.3.16] we know that $\xi_k \in \pi_u \partial f(u_k, v_k)$ (where π_u denotes projection on the u space). Therefore we can find a sequence $\{\eta_k\}$ such that $\{(\xi_k, \eta_k)\} \in \partial f(u_k, v_k)$. By the local boundedness and upper-semicontinuity of the generalized gradient of a locally Lipschitz function, we then see that we must have (renumbering if necessary) $\{(\xi_k, \eta_k)\} \to \{(\bar{\xi}, \bar{\eta})\} \in \partial f(\bar{u}, \bar{v})$. But then, by the convexity assumption and [8, Proposition 2.5.3], we conclude that $\bar{\xi} \in \partial_u f(\bar{u}, \bar{v})$. □

We can now investigate the convergence properties of Algorithm 3.

Theorem 4.3 *Assume that Assumptions* **A1–A3** *hold. Let $\{\boldsymbol{x}_k\}$ be the sequence generated by Algorithm 3 and let $\bar{\boldsymbol{x}}$ be a cluster point of this sequence. Then $\bar{\boldsymbol{x}}$ is a Nash equilibrium of Problem (15).*

Proof By Lemma 4.1, we have that for every k and every ν $x_k^\nu \in X_\nu(\boldsymbol{x}_{k,\nu}^{-\nu})$. By this and by the definition of x_{k+1}^ν in Step 3, we then have

$$\theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) \le \theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \tau \|x_{k+1}^\nu - x_k^\nu\|^2, \quad \forall k, \nu. \tag{17}$$

By the definition of Generalized Potential Game, this relationship implies that

$$P(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - P(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) \ge \sigma(\theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})) \ge 0, \quad \forall k, \nu. \tag{18}$$

Noting that $\boldsymbol{x}_{k,\nu} = (x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})$ and $\boldsymbol{x}_{k,\nu+1} = (x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})$, we get from (18)

$$P(\boldsymbol{x}_{k,\nu+1}) \le P(\boldsymbol{x}_{k,\nu}). \tag{19}$$

From (19), recalling that $\boldsymbol{x}_k = \boldsymbol{x}_{k,1}$, and $\boldsymbol{x}_{k+1} = \boldsymbol{x}_{k,N+1}$, we get

$$P(\boldsymbol{x}_{k+1}) = P(\boldsymbol{x}_{k,N+1}) \le \cdots \le P(\boldsymbol{x}_{k,\nu}) \le \cdots \le P(\boldsymbol{x}_{k,1}) = P(\boldsymbol{x}_k). \tag{20}$$

Let $K \subseteq \{0, 1, \ldots\}$ be an infinite subset of indices such that $\lim_{k\to\infty, k\in K} \boldsymbol{x}_k = \bar{\boldsymbol{x}}$. By the continuity of P and by (20) it follows that the full sequence $\{P(\boldsymbol{x}_k)\}$ is convergent to a finite value $\bar{P}$, and, therefore, again by (20) it also follows that

$$\lim_{k\to\infty} P(\boldsymbol{x}_{k,\nu}) = \bar{P}, \quad \forall \nu. \tag{21}$$

In turn, taking into account (18), this implies

$$\lim_{k\to\infty} \sigma(\theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})) = 0, \tag{22}$$

and hence by definition of forcing function

$$\lim_{k\to\infty}(\theta_\nu(x_k^\nu,\boldsymbol{x}_{k,\nu}^{-\nu})-\theta_\nu(x_{k+1}^\nu,\boldsymbol{x}_{k,\nu}^{-\nu}))=0, \tag{23}$$

that combined with (17) gives

$$\lim_{k\to\infty}\|x_{k+1}^\nu-x_k^\nu\|=0. \tag{24}$$

By (24) and the definition of $\boldsymbol{x}_{k,\nu}$ in Step 2, we then also have

$$\lim_{k\to\infty,k\in K}\boldsymbol{x}_{k,\nu}=\bar{\boldsymbol{x}},\quad \forall\nu. \tag{25}$$

By Lemma 4.1 and by the closure of $X_\nu(\cdot)$ for all ν, we have $\bar{x}^\nu\in X_\nu(\bar{\boldsymbol{x}}^{-\nu})$, for all ν. We prove that

$$\theta_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu})\le\theta_\nu(x^\nu,\bar{\boldsymbol{x}}^{-\nu}),\quad \forall x^\nu\in X_\nu(\bar{\boldsymbol{x}}^{-\nu}). \tag{26}$$

By contradiction assume that there exists a ν and a vector $\bar{y}^\nu\in X_\nu(\bar{\boldsymbol{x}}^{-\nu})$ such that

$$\theta_\nu(\bar{y}^\nu,\bar{\boldsymbol{x}}^{-\nu})<\theta_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu}).$$

By Assumption **A3**, the problem of the ν-th player is convex. Therefore θ_ν is directionally differentiable at $(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu})$ and, with $d^\nu\equiv(\bar{y}^\nu-\bar{x}^\nu)$, we can write

$$\begin{aligned}\theta'_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu};d^\nu)&=\max_{\xi\in\partial_{x^\nu}\theta_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu})}\xi^Td^\nu\\&=\inf_{\lambda>0}\frac{\theta_\nu(\bar{x}^\nu+\lambda d^\nu,\bar{\boldsymbol{x}}^{-\nu})-\theta_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu})}{\lambda}\\&\le\frac{\theta_\nu(\bar{y}^\nu,\bar{\boldsymbol{x}}^{-\nu})-\theta_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu})}{1}<0.\end{aligned} \tag{27}$$

From the inner-semicontinuity of $X_\nu(\cdot)$ and from (25), it follows that there exists a sequence $\{y_k^\nu\}$ such that $y_k^\nu\in X(\boldsymbol{x}_{k,\nu}^{-\nu})$ and $\lim_{k\to\infty}y_k^\nu=\bar{y}^\nu$. Let us set $\Theta(x^\nu,\boldsymbol{x}^{-\nu},z)\equiv\theta(x^\nu,\boldsymbol{x}^{-\nu})+\frac{1}{2}\|x^\nu-z\|^2$. Note that $\Theta'(x^\nu,\boldsymbol{x}^{-\nu},z;d^\nu)=\theta'_\nu(x^\nu,\boldsymbol{x}^{-\nu};d^\nu)+(x^\nu-z)^Td^\nu$, so that recalling the definition of x_{k+1}^ν at Step 2, the optimality conditions for a convex problem and elementary properties of convex functions, we can write, for some suitable $\xi^k\in\partial_{x^\nu}\theta(x_{k+1}^\nu,\boldsymbol{x}_{k,\nu}^{-\nu})$,

$$\begin{aligned}&\Theta'(x_{k+1}^\nu,\boldsymbol{x}_{k,\nu}^{-\nu},x_k^\nu;(y_k^\nu-x_{k+1}^\nu))\\&\quad=\theta'(x_{k+1}^\nu,\boldsymbol{x}_{k,\nu}^{-\nu};y_k^\nu-x_{k+1}^\nu)+(x_{k+1}^\nu-x_k^\nu)^T(y_k^\nu-x_{k+1}^\nu)\\&\quad=(\xi^k)^T(y_k^\nu-x_{k+1}^\nu)+(x_{k+1}^\nu-x_k^\nu)^T(y_k^\nu-x_{k+1}^\nu)\ge 0.\end{aligned}$$

Passing to the limit for $k\to\infty$, $k\in K$, using Proposition 4.2, (24) and (25), we get, for some suitable $\bar{\xi}\in\partial_{x^\nu}\theta(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu})$,

$$0\le\bar{\xi}^T(\bar{y}^\nu-\bar{x}^\nu)\le\theta'_\nu(\bar{x}^\nu,\bar{\boldsymbol{x}}^{-\nu};(\bar{y}^\nu-\bar{x}^\nu)).$$

This contradicts (27) and concludes the proof. □

The previous theorem leaves open the question whether the sequence generated by Algorithm 3 admits a cluster point. Because of Lemma 4.1 this would obviously be true if the feasible set $\{\boldsymbol{x} \in D_1 \times \cdots \times D_N : \boldsymbol{x} \in \boldsymbol{X}\}$ is bounded. A more general assumption could be that the potential function be coercive, but we do not go into these details here.

Remark 4.1 For sake of simplicity, in Algorithm 3, we took τ fixed and the same for every player. It is however easy to see that all the results of this section go through if each player uses a different τ_ν and if these τ_ν vary from iteration to iteration, provided they are all bounded away from zero and from above.

Remark 4.2 By way of example we consider the three players potential Nash game described in Sect. 3.1. Assumptions **A1**–**A3** hold so that Algorithm GS-C can be applied and the convergence is guaranteed. We take as starting point the point $(-1-\epsilon, 1+\frac{1}{2}\epsilon, -1-\frac{1}{4}\epsilon)$ with $\varepsilon = 0.1$. At each iteration k, at Step 3 of Algorithm GS-C subproblem (16) is a univariate and convex problem that can actually be solved analytically. We set $\tau = 0.5$, and the algorithm converged in 8 iterations to the Nash equilibrium point $(-10, -10, -10)$ (we terminated the algorithm when the norm of the violation of the KKT conditions of all the players was below 10^{-6}).

5 The case of N players without convexity

In this section we drop Assumption **A3** and show that we can still obtain useful convergence results using decomposition, even when the players' problem are not convex. The "price" we have to pay to get these results is that non convex subproblems have to be solved by the players at each iteration, and this can obviously be difficult in general. However, the problems the players have to solve in the decomposition scheme are often simple (if compared to the global solution of the optimization problem (7)); for example, this is the case for the application discussed in Sect. 3.4.

The key point in order to achieve convergence in the non convex case is a suitable law for controlling the parameter τ in Algorithm 2 that forces it to zero in a controlled way. The resulting algorithm is the following.

A first observation concerns the stopping criterion in Step 1. Since we dropped any convexity assumption, this step is not totally straightforward in principle. However, since a solution $\bar{\boldsymbol{x}}$ of the game is a fixed point of the iteration in Step 2 (in the sense that if $\boldsymbol{x}_k = \bar{x}$ then $\boldsymbol{x}_{k+1} = \bar{x}$), then a suitable stopping criterion is $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\| \le \text{tol}$, where tol is a prefixed positive tolerance.

The question also arises of whether subproblems (28) do have a solution, so that the algorithm is well defined. These subproblems will obviously have solutions under a host of standard assumptions. For example it would be enough to assume that the set $X_\nu(\boldsymbol{x}^{-\nu})$ is non empty and compact (which in turn is true if either $\boldsymbol{X}$ or D_ν are compact). In the sequel, to avoid such irrelevant hypotheses, we simply require that:

A4 For all ν and for all k, subproblem (28) has a solution.

The main difference of Algorithm 4 with respect to Algorithm 3 is obviously Step 3. We will see that the updating rule (29) forces the regularization parameter

Algorithm 4 Regularized Gauss-Seidel—NonConvex Subproblems

`(S.0)`: Choose any feasible starting point $\boldsymbol{x}_0 = (x_0^1, \dots, x_0^N)$, a positive regularization parameter $\tau_0 > 0$ and set $k := 0$.

`(S.1)`: If $\boldsymbol{x}_k$ satisfies a suitable termination criterion: `STOP`.
Otherwise set $\boldsymbol{x}_{k,1} = \boldsymbol{x}_k$.

`(S.2)`: `for` $\nu = 1, \dots, N$, compute a (global) solution x_{k+1}^ν of

$$\begin{aligned} \min_{x^\nu} \quad & \theta_\nu(x_{k+1}^1, \dots, x_{k+1}^{\nu-1}, x^\nu, x_k^{\nu+1}, \dots, x_k^N) + \tau_k \|x^\nu - x_k^\nu\|^2 \\ \text{s.t.} \quad & x^\nu \in X_\nu(x_{k+1}^1, \dots, x_{k+1}^{\nu-1}, x_k^{\nu+1}, \dots, x_k^N). \end{aligned} \tag{28}$$

Set $\boldsymbol{x}_{k,\nu+1} = (x_{k+1}^1, \dots, x_{k+1}^\nu, x_k^{\nu+1}, \dots, x_k^N)$.
`end`

`(S.3)`: Set

$$\tau_{k+1} = \max\left\{\min\left[\tau_k, \max_{\nu=1,\dots,N}\{\|x_{k+1}^\nu - x_k^\nu\|\}\right], 0.1\tau_k\right\}, \tag{29}$$

$\boldsymbol{x}_{k+1} := (x_{k+1}^1, \dots, x_{k+1}^N)$, $k \leftarrow k+1$, and go to `(S.1)`.

τ_k to zero, without actually ever letting it be zero, and therefore τ_k is reduced an infinite number of times. Let K be the subsequence of iterations where τ_k is reduced. In Theorem 5.2 below we show that every limit point of $\{\boldsymbol{x}_k\}_K$ is a Nash equilibrium. To this end we preliminarily observe that Lemma 4.1 is easily seen to be still valid in the current setting under Assumption **A4**.

Lemma 5.1 *Suppose Assumption* **A4** *holds. If there exists a cluster point of the sequence* $\{\boldsymbol{x}_k\}$, *then*

(i) *we have*

$$\lim_{k\to\infty} \tau_k = 0 \tag{30}$$

(ii) *there exists an infinite index set* K *of iterations such that*

$$\tau_{k+1} < \tau_k. \tag{31}$$

Proof From the instructions at Step 3 we have

$$\tau_{k+1} \le \tau_k.$$

In order to prove point (i), we assume by contradiction that

$$\tau_k \ge \bar{\tau} > 0, \quad \forall k. \tag{32}$$

By Lemma 4.1, we have that for every k and for every ν $x_k^\nu \in X_\nu(\boldsymbol{x}_{k,\nu}^{-\nu})$. By this and by the definition of x_{k+1}^ν in Step 2, we then have

$$\theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) \le \theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \bar{\tau}\|x_{k+1}^\nu - x_k^\nu\|^2, \quad \forall k, \nu. \tag{33}$$

By definition of Generalized Potential Game this relationship implies that

$$P(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - P(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) \geq \sigma(\theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})) \geq 0, \quad \forall k, \nu. \tag{34}$$

Noting that $\boldsymbol{x}_{k,\nu} = (x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})$ and $\boldsymbol{x}_{k,\nu+1} = (x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})$, we can rewrite (34) as

$$P(\boldsymbol{x}_{k,\nu+1}) \leq P(\boldsymbol{x}_{k,\nu}), \quad \forall \nu. \tag{35}$$

From (35), recalling that $\boldsymbol{x}_k = \boldsymbol{x}_{k,1}$, and $\boldsymbol{x}_{k+1} = \boldsymbol{x}_{k,N+1}$, we get

$$P(\boldsymbol{x}_{k+1}) = P(\boldsymbol{x}_{k,N+1}) \leq \cdots \leq P(\boldsymbol{x}_{k,\nu}) \leq \cdots \leq P(\boldsymbol{x}_{k,1}) = P(\boldsymbol{x}_k). \tag{36}$$

Let $\hat{K} \subseteq \{0, 1, \ldots\}$ be an infinite subset of indices such that

$$\lim_{k\to\infty, k\in\hat{K}} \boldsymbol{x}_k = \bar{\boldsymbol{x}}.$$

By the continuity of P and by (36) it follows that the full sequence $\{P(\boldsymbol{x}_k)\}$ is convergent to a finite value $\bar{P}$, and, therefore, again by (36) it also follows that

$$\lim_{k\to\infty} P(\boldsymbol{x}_{k,\nu}) = \bar{P}, \quad \forall \nu. \tag{37}$$

In turn, taking into account (34), this implies

$$\lim_{k\to\infty} \sigma(\theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})) = 0, \tag{38}$$

and hence by definition of forcing function

$$\lim_{k\to\infty} (\theta_\nu(x_k^\nu, \boldsymbol{x}_{k,\nu}^{-\nu}) - \theta_\nu(x_{k+1}^\nu, \boldsymbol{x}_{k,\nu}^{-\nu})) = 0, \tag{39}$$

that combined with (33) gives

$$\lim_{k\to\infty} \|x_{k+1}^\nu - x_k^\nu\| = 0. \tag{40}$$

From (40) we get

$$\max_{1,\ldots,N} \{\|x_{k+1}^1 - x_k^1\|, \ldots, \|x_{k+1}^\nu - x_k^\nu\|, \ldots, \|x_{k+1}^N - x_k^N\|\} < \bar{\tau}$$

for k sufficiently large, which implies, together with (29), that $\tau_{k+1} < \bar{\tau}$, and this contradicts (32).

Point (ii) follows from point (i) and the updating rule (29). □

Theorem 5.2 *Suppose that Assumptions* **A1**, **A2** *and* **A4** *hold. Assume that there exists a cluster point of the sequence* $\{\boldsymbol{x}_k\}$. *Let* K *be the infinite subset of iterations defined at point* (*ii*) *of Lemma* 5.1. *Then any cluster point of the subsequence* $\{\boldsymbol{x}_k\}_K$ *is a Nash equilibrium of Problem* (15).

Proof Let $K \subseteq \{0, 1, \ldots\}$ be the infinite subset of iterations defined at point (ii) of Lemma 5.1 where

$$\tau_{k+1} < \tau_k.$$

Then for all $k \in K$ we have

$$\max_{1,\ldots,N} \{\|x^1_{k+1} - x^1_k\|, \ldots, \|x^\nu_{k+1} - x^\nu_k\|, \ldots, \|x^N_{k+1} - x^N_k\|\} < \tau_k.$$

From the above inequality and from point (i) of Lemma 5.1 we obtain

$$\lim_{k\to\infty, k\in K} \|x^\nu_{k+1} - x^\nu_k\| = 0, \quad \forall \nu. \tag{41}$$

Let $\bar{\boldsymbol{x}}$ be any cluster point of the subsequence $\{\boldsymbol{x}_k\}_K$. By (41) and the definition of $\boldsymbol{x}_{k,\nu}$ in Step 3, we then also have

$$\lim_{k\to\infty, k\in K} \boldsymbol{x}_{k,\nu} = \bar{\boldsymbol{x}}, \quad \forall \nu. \tag{42}$$

By Lemma 4.1 and by the closure of $X_\nu(\cdot)$ for all ν, we have $\bar{x}^\nu \in X_\nu(\bar{\boldsymbol{x}}^{-\nu})$, for all ν. We prove that

$$\theta_\nu(\bar{x}^\nu) \le \theta_\nu(x^\nu), \quad x^\nu \in X_\nu(\bar{\boldsymbol{x}}^{-\nu}). \tag{43}$$

By contradiction assume that there exists a ν and a vector $\bar{y}^\nu \in X_\nu(\bar{\boldsymbol{x}}^{-\nu})$ such that

$$\theta_\nu(\bar{y}^\nu) < \theta_\nu(\bar{x}^\nu). \tag{44}$$

From the inner-semicontinuity of $X_\nu(\cdot)$ and from (42), it follows that there exists a sequence $\{y^\nu_k\}$ such that $y^\nu_k \in X(\boldsymbol{x}^{-\nu}_{k,\nu})$ and

$$\lim_{k\to\infty, k\in K} y^\nu_k = \bar{y}^\nu.$$

Then, recalling the definition of x^ν_{k+1} at Step 3 we can write

$$\theta_\nu(x^\nu_{k+1}) + \tau_k \|x^\nu_{k+1} - x^\nu_k\|^2 \le \theta_\nu(y^\nu_k) + \tau_k \|y^\nu_k - x^\nu_k\|^2.$$

Passing to the limit for $k \to \infty$, $k \in K$, recalling point (i) of Lemma 5.1 and (41) we get a contradiction to (44) and this concludes the proof. □

Note that the previous result is based on the existence of a limit point of $\{\boldsymbol{x}_k\}$ which, however, does not ensure that limit points of $\{\boldsymbol{x}_k\}_K$ exist. What it claims is that if limit points of this latter sequence exist, then these are solution. In order to be sure that such limit points exist it is sufficient to require that $D_\nu \cap \boldsymbol{X}$ be bounded for all ν.

We believe that the results of this theorem are interesting, since they are obtained without making any of the standard convexity assumptions usually encountered when studying (decomposition) algorithms for the solution of games. Furthermore, the theorem also makes clear that the regularization term in Algorithm 4 plays a role rather different from the traditional one, since it is applied to a fully non convex problem.

A final simple example might be useful to prove that in the non convex setting of this section we really need the parameter τ_k to decrease to zero and can not keep it fixed, as in the previous section. To this end we show that given any *fixed* positive parameter τ, we can find a problem for which Algorithm 3 does not converge to Nash equilibria. So, let a *fixed*, positive τ be given. Consider a game with two players, $N = 2$. Assume that each player controls only one variable, $n_1 = 1$ and $n_2 = 2$, and let the players' problems be

$$\begin{array}{llll} \min_x & -\frac{\tau}{2}x^2 + xy, & \min_y & y^2 + xy, \\ & 0 \le x \le 1, & & 0 \le y, \\ & y \le x + 1, & & y \le x + 1. \end{array}$$

It is obvious that this problem has a unique solution given by $(\bar{x}, \bar{y}) = (1, 0)$. If we set $(x_0, y_0) = (0, 0)$ and apply one step of the regularization procedure, we get that x_1 will be the solution of the regularized problem

$$\begin{array}{ll} \min_x & -\frac{\tau}{2}x^2 + 0x + \tau(x-0)^2, \\ & 0 \le x \le 1, \\ & 0 \le x + 1, \end{array}$$

and it is easy to see that $x_1 = 0$. Consequently, y_1 will be the solution of the regularized problem

$$\begin{array}{ll} \min_y & y^2 + 0y + \tau(y-0)^2, \\ & 0 \le y, \\ & y \le 0 + 1, \end{array}$$

and even in this case it is immediate to check that $y_1 = 0$. Therefore $(x_1, y_1) = (x_0, y_0) = (0, 0)$, and it is clear that the application of Algorithm 3 will generate a sequence $\{(x_k, y_k)\}$ with $(x_k, y_k) = (0, 0)$ for every k. This sequence converges to $(0, 0)$ which is not the solution of the problem.

6 The case of two players

The case of two players is particularly interesting for two reasons. First, it often occurs in practice and represents the "antagonistic" behavior in its purest form. Second, in the case of two players, stronger results can be obtained. In fact, we can somewhat enlarge the class of problems for which we can show convergence by relaxing condition (b) in the Definition 2.1 of Generalized Potential Game. But the most striking fact about the 2 players case is that there is no need of regularization.

Let the following GNEP with two players be given:[2]

$$\begin{array}{llcll} \text{minimize}_x & \theta_1(x,y) & \multirow{2}{*}{and} & \text{minimize}_y & \theta_2(x,y) \\ \text{subject to} & x \in X(y) & & \text{subject to} & y \in Y(x) \end{array} \tag{45}$$

where $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$, $\theta_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$, $\theta_2 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}$ and the following two conditions are satisfied for a suitable potential function P:

(a)

$$X(y) = \{x \in D_x : (x,y) \in \boldsymbol{X}\}, \quad Y(x) = \{y \in D_y : (x,y) \in \boldsymbol{X}\},$$
$$\text{with } D_x \subseteq \mathbb{R}^{n_1},\ D_y \subseteq \mathbb{R}^{n_2} \text{ and } \boldsymbol{X} \subseteq \mathbb{R}^{n_1+n_2} \text{ closed sets,}$$

(b)

$$\theta_1(\bar{x},y) - \theta_1(\tilde{x},y) \geq (>)0 \quad \Longrightarrow \quad P(\bar{x},y) - P(\tilde{x},y) \geq (>)0,$$
$$\theta_2(x,\bar{y}) - \theta_2(x,\tilde{y}) \geq (>)0, \quad \Longrightarrow \quad P(x,\bar{y}) - P(x,\tilde{y}) \geq (>)0.$$

In the next theorem we show that we can apply Algorithm 2 taking $\tau_k = 0$ for every k. This means that we are actually applying Algorithm 1, and this is the only case in which the latter algorithm actually works. For sake of clarity we report the algorithm for the present setting.

Algorithm 5 Gauss-Seidel—Two Players

`(S.0)`: Choose a feasible starting point $x_0 \in \mathbb{R}^{n_1}$, $y_0 \in \mathbb{R}^{n_2}$ and set $k := 0$.
`(S.1)`: If (x_k, y_k) satisfies a suitable termination criterion: `STOP`.
`(S.2)`: Let x_{k+1} be a (global) solution of

$$\begin{aligned} &\min \theta_1(x, y_k) \\ &x \in X(y_k). \end{aligned} \tag{46}$$

`(S.3)`: Let y_{k+1} be a (global) solution of

$$\begin{aligned} &\min \theta_2(x_{k+1}, y) \\ &y \in Y(x_{k+1}). \end{aligned} \tag{47}$$

`(S.4)`: Set $k \leftarrow k+1$, and go to `(S.1)`.

As usual, we assume that subproblems (46) and (47) have a solution at each step, and call this Assumption **A4**. Furthermore note that, by an immediate extension of Lemma 4.1 we have that

$$x_k \in X(y_k), \quad \forall k \geq 1, \tag{48}$$

[2]For sake of clarity in this section we deviate from the general notation adopted and indicate the variables of the first player with x and those of the second with y.

$$y_k \in Y(x_{k+1}), \quad \forall k \geq 1. \tag{49}$$

With this preparation, we can now prove the following theorem.

Theorem 6.1 *Assume that Assumptions* **A1**, **A2**, *and* **A4** *hold. Let* $\{(x_k, y_k)\}$ *be the sequence generated by the Algorithm* 5. *Let* $(\bar{x}, \bar{y})$ *be a cluster point of* $\{(x_k, y_k)\}$. *Then* $(\bar{x}, \bar{y})$ *is a Nash equilibrium of Problem* (45).

Proof By steps 2 and 3, for every k we have

$$\theta_1(x_{k+1}, y_k) \leq \theta_1(x, y_k), \quad \forall x \in X(y_k), \tag{50}$$

$$\theta_2(x_{k+1}, y_{k+1}) \leq \theta_2(x_{k+1}, y), \quad \forall y \in Y(x_{k+1}). \tag{51}$$

Using (48), (49), (50), (51) we obtain

$$\theta_1(x_{k+1}, y_k) \leq \theta_1(x_k, y_k), \tag{52}$$

$$\theta_2(x_{k+1}, y_{k+1}) \leq \theta_2(x_{k+1}, y_k). \tag{53}$$

Since condition (b) in this section holds, we then get

$$P(x_k, y_k) - P(x_{k+1}, y_k) \geq 0, \tag{54}$$

$$P(x_{k+1}, y_k) - P(x_{k+1}, y_{k+1}) \geq 0, \tag{55}$$

that imply

$$P(x_{k+1}, y_{k+1}) \leq P(x_{k+1}, y_k) \leq P(x_k, y_k). \tag{56}$$

Let $K \subseteq \{0, 1, \ldots\}$ be an infinite subset of indices such that

$$\lim_{k\to\infty, k\in K} (x_k, y_k) = (\bar{x}, \bar{y}).$$

By the continuity of P and by (56), we have

$$\lim_{k\to\infty} P(x_k, y_k) = \lim_{k\to\infty} P(x_{k+1}, y_k) = \bar{P}. \tag{57}$$

Since $X(\cdot)$ and $Y(\cdot)$ are closed, we have that $\bar{x} \in X(\bar{y})$ and $\bar{y} \in Y(\bar{x})$. We prove that

$$\theta_1(\bar{x}, \bar{y}) \leq \theta_1(x, \bar{y}), \quad \forall x \in X(\bar{y}), \tag{58}$$

$$\theta_2(\bar{x}, \bar{y}) \leq \theta_2(\bar{x}, y), \quad \forall y \in Y(\bar{x}). \tag{59}$$

By contradiction assume first that (58) does not hold, i.e., that there exists a vector $\hat{x} \in X(\bar{y})$ such that

$$\theta_1(\hat{x}, \bar{y}) < \theta_1(\bar{x}, \bar{y}), \tag{60}$$

that implies

$$P(\bar{x},\bar{y}) - P(\hat{x},\bar{y}) > 0. \tag{61}$$

From the inner-semicontinuity of $X(\cdot)$ it follows that there exists a sequence $\{\hat{x}_k\}_K$ such that $\hat{x}_k \in X(y_k)$ for all $k \in K$ and $\lim_{k\to\infty, k\in K} \hat{x}_k = \hat{x}$. From (50) we get

$$\theta_1(x_{k+1}, y_k) \leq \theta_1(\hat{x}_k, y_k),$$

that implies

$$P(\hat{x}_k, y_k) - P(x_{k+1}, y_k) \geq 0. \tag{62}$$

The continuity of P and (57) imply

$$P(\bar{x},\bar{y}) = \bar{P}. \tag{63}$$

Taking the limits for $k \to \infty$, $k \in K$, taking into account (57) and (63), we obtain

$$P(\bar{x},\bar{y}) \leq P(\hat{x},\bar{y}),$$

which contradicts (61).

Now, again by contradiction assume that (59) does not hold, i.e., that there exists a vector $\hat{y} \in Y(\bar{x})$ such that

$$\theta_2(\bar{x},\hat{y}) < \theta_2(\bar{x},\bar{y}), \tag{64}$$

and hence

$$P(\bar{x},\bar{y}) - P(\bar{x},\hat{y}) > 0. \tag{65}$$

From the inner-semicontinuity of $Y(\cdot)$ it follows that there exists a sequence $\{\hat{y}_k\}_K$ such that $\hat{y}_k \in Y(x_k)$ for all $k \in K$ and $\lim_{k\to\infty, k\in K} \hat{y}_k = \hat{y}$. From (51) we get

$$\theta_2(x_k, y_k) \leq \theta_2(x_k, \hat{y}_k),$$

and hence

$$P(x_k, \hat{y}_k) - P(x_k, y_k) \geq 0,$$

Taking the limits for $k \to \infty$, $k \in K$, recalling (63), we obtain

$$P(\bar{x},\bar{y}) \leq P(\bar{x},\hat{y}),$$

which contradicts (65). Thus (58) and (59) show that $(\bar{x},\bar{y})$ is a Nash equilibrium of (45). □

This theorem could seem to suggest that our analysis in previous sections was not deep enough and that maybe it could be possible to eliminate the regularization terms also in the case of games with more than two players. Unfortunately this is not so, as shown by the example in Sect. 3.1. The case of two players is really peculiar.

Remark 6.1 Antipin devoted some papers (see [3] and references therein) to the study of convergence of regularized projection-like methods for the solution of two-person games. His algorithms are close in spirit to those considered in this paper, and his (rather complex) assumptions also appear to be strongly related to that of potential game we used here (although, among other things, he certainly assumes convexity of the feasible regions and of the objective functions of each player for a fixed strategy of the opponent). It may be interesting to note then that we show that regularization is not really needed in the case of two players only. On the other hand, Antipin is able to show convergence of the whole sequence to a single solution.

Acknowledgements We are very grateful to Diethard Klatte and to Gesualdo Scutari for their invaluable help with many aspects of this paper. We are also grateful to Christian Kanzow who read a preliminary version of the paper and gave useful suggestions.

References

1. Altman, E., Boulogne, T., El-Azouzi, R., Jiménez, T., Wynter, L.: A survey on networking games in telecommunications. Comput. Oper. Res. **33**, 286–311 (2006)
2. Altman, E., Wynter, L.: Equilibrium, games and pricing in transportation and telecommunication Networks. Netw. Spat. Econ. **4**, 7–21 (2004)
3. Antipin, A.: Extra-proximal methods for solving two-person nonzero-sum games. Math. Program. B (2007). doi:10.1007/s10107-007-0152-2. Published online
4. Aussel, D., Dutta, J.: Generalized Nash equilibrium problem, variational inequality and quasiconvexity. Oper. Res. Lett. **36**, 461–464 (2008)
5. Bank, B., Guddat, J., Klatte, D., Kummer, B., Tammer, K.: Non-linear Parametric Optimization. Akademie-Verlag, Berlin (1982)
6. Berman, A., Plemmons, R.J.: Nonnegative Matrices in the Mathematical Sciences. SIAM, Philadelphia (1994)
7. Chung, M., Hande, P., Lan, T., Tan, C.W.: Power Control in Wireless Cellular Networks. Now Publishers Inc., Boston (2008)
8. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)
9. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
10. Ermoliev, Yu.M., Flåm, S.D.: Repeated play of potential games. Cybern. Syst. Anal. **38**, 355–367 (2002)
11. Facchinei, F., Kanzow, C.: Generalized Nash equilibrium problems. 4OR **5**, 173–210 (2007)
12. Facchinei, F., Pang, J.-S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, Berlin (2003)
13. Facchinei, F., Pang, J.-S.: Nash equilibria: the variational approach. In: Palomar, D.P., Eldar, J.C. (eds.) Convex Optimization in Signal Processing and Communications, pp. 443–493. Cambridge University Press, Cambridge (2010)
14. Foschini, G.J., Miljanic, Z.: A simple distributed autonomous power control algorithm and its convergence. IEEE Trans. Veh. Technol. **42**, 641–646 (1993)
15. Grippo, L., Sciandrone, M.: On the convergence of the block nonlinear Gauss-Seidel method under convex constraints. Oper. Res. Lett. **26**, 127–136 (2000)
16. Matioli, L.C., Sosa, W., Yuan, J.: A numerical algorithm for finding solutions of a generalized Nash equilibrium problem. Imca Technical Report RI-04-2009, Lima Perù (2009)
17. Monderer, D., Shapley, L.S.: Potential Games. Games Econ. Behav. **14**, 124–143 (1996)
18. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: Algorithmic Game Theory. Cambridge University Press, Cambridge (2007)
19. O'Neill, D., D. Julian, D., Boyd, S.: Seeking Foschini's genie: optimal rates and powers in wireless networks. IEEE Trans. Veh. Technol. (accepted for publication)
20. Powell, M.J.D.: On search directions for minimization algorithms. Math. Program. **4**, 193–201 (1973)
21. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)

22. Rosen, J.B.: Existence and uniqueness of equilibrium points for concave n-person games. Econometrica **33**, 520–534 (1965)
23. Scutari, G., Palomar, D., Facchinei, F., Pang, J.-S.: Flexible design of cognitive radio wireless systems: from game theory to variational inequality theory. IEEE Signal Process. Mag. **26**, 107–123 (2009)
24. Tse, D., Viswanath, P.: Fundamental of Wireless Communications. Cambridge University Press, Cambridge (2005)
25. Viterbi, A.J.: CDMA: Principles of Spread Spectrum Communication. Addison-Wesley, Reading (1995)
26. Yates, R.D.: A framework for uplink power control in cellular radio systems. IEEE J. Sel. Areas Commun. **52**, 1341–1347 (1995)